

[고급 비즈니스 어널리틱스]

# 전자상거래 데이터 분석: 지역별 배송 지연 분석 및 솔루션 제시

1조  
조창연, 김선엽, 임은석

# 목차

1. Data overview

2. Data Story Catch

3. Data 전처리

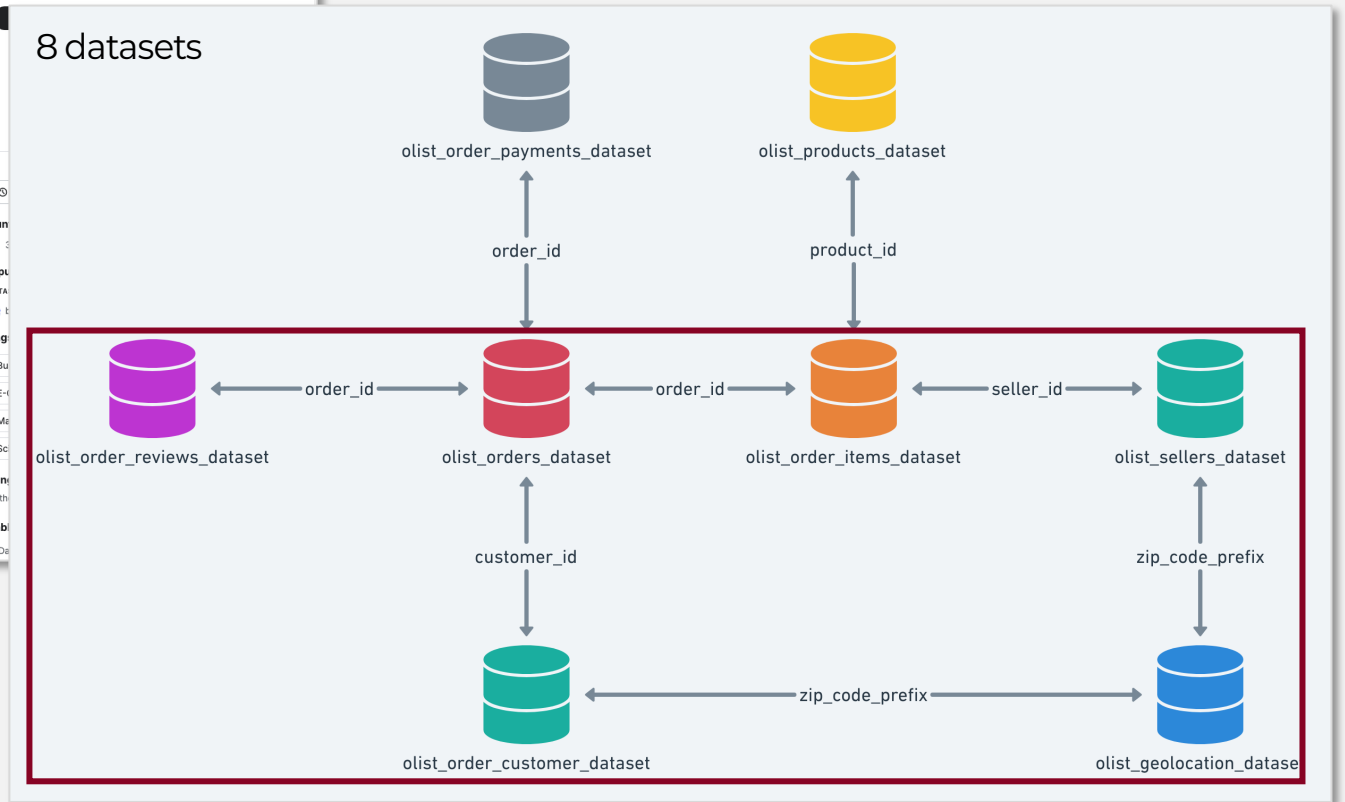
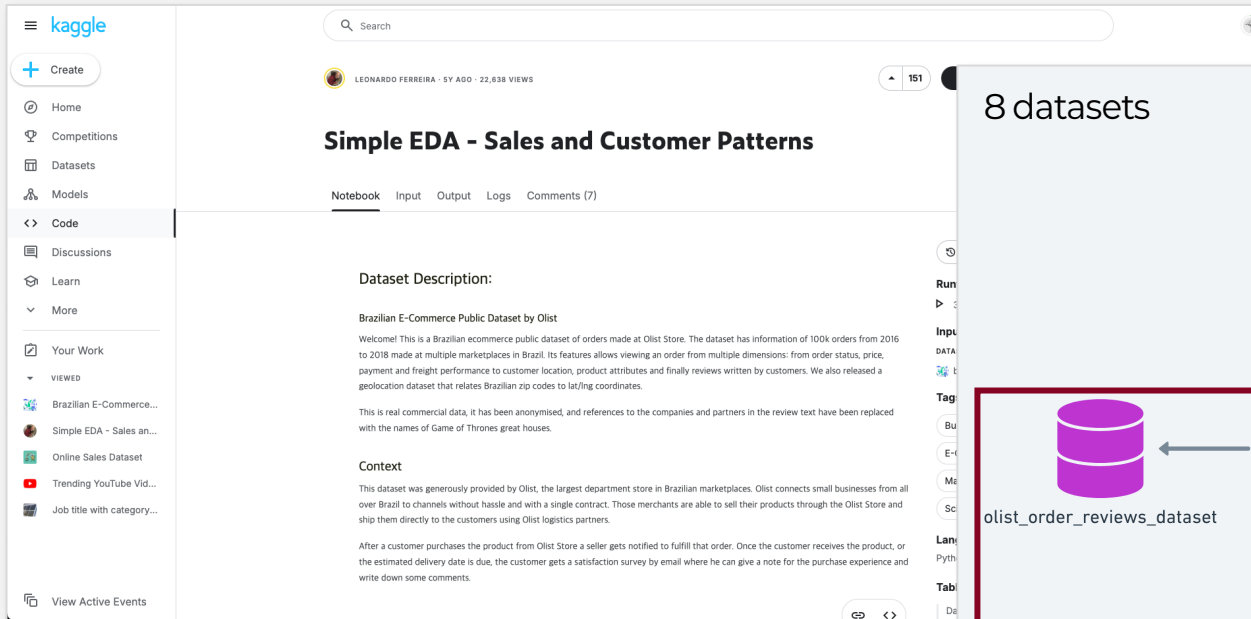
4. 데이터 시각화 및 문제 도출

5. 해결 전략 제시

# 1. Data Overview

# Simple EDA - Sales and Customer Patterns

Kaggle에서 제공하는 브라질 e-커머스 플랫폼 데이터  
고객 및 상품 정보뿐만 아니라, 배송과 관련된 상세한 정보 포함



## **2. Data Story Catch**

# Data Story Catch

SP 지역에서 발생하는 배송 지연 문제와 그로 인한 주문량 감소를 해결하기 위해 작성

## Step 1: Problem

상품의 배송 기간이 길어질수록 고객만족도와 주문량이 감소할 것이다.

## Step 2: Statement → Question

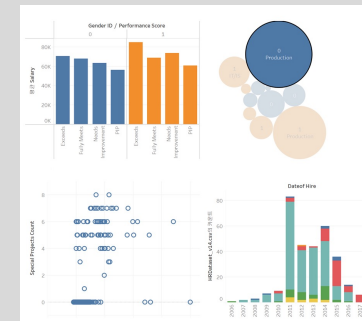
배송 시간이 고객 만족도와 매출에 어떤 영향을 미치는가?

어떤 도시들이 배송 시간에 문제가 있을까?

## Step 3: Find Data

[Kaggle의 브라질 e-커머스 데이터셋 활용](#)

## Step 4: Visualize



### 3. Data 전처리

## 4. 데이터 시각화 및 문제 도출



## 주문부터 배송 완료까지의 단계별 평균 시간

배송의 각 구간별 평균 소요 시간을 계산하여 분석  
각 구간별로 계산한 결과, 총 평균 기간 13일

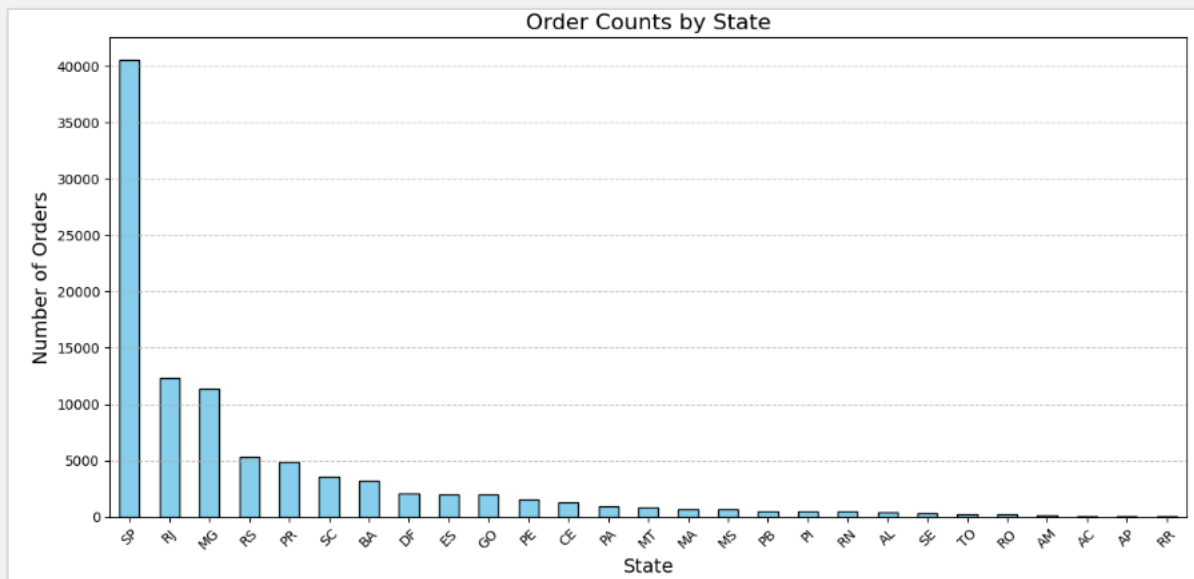


## 주문 지역 집중화에 따른 문제 인식

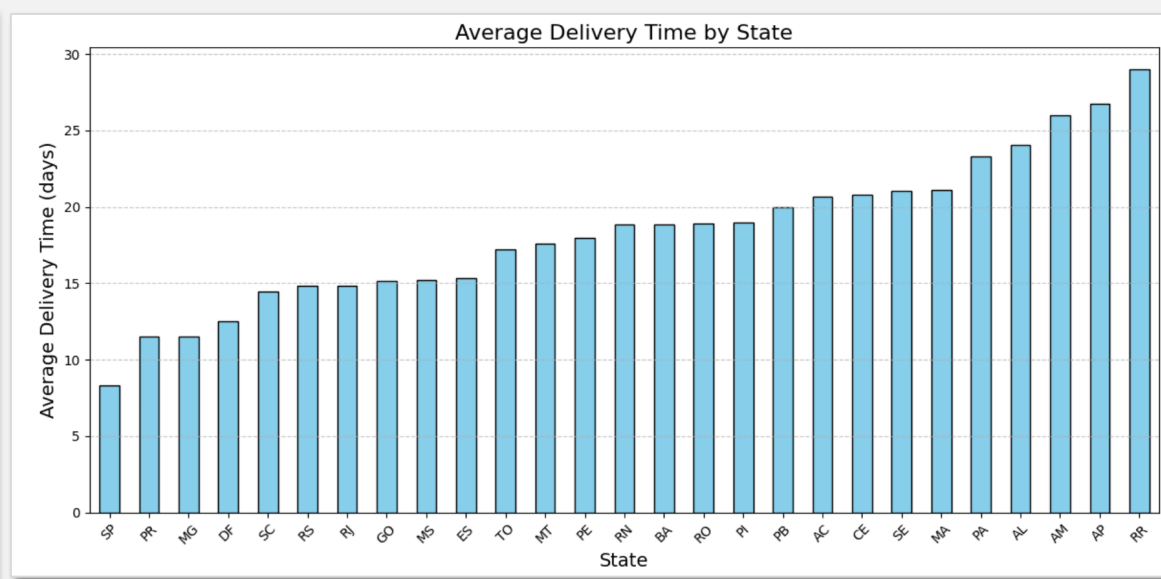
소비자의 거주 지역을 기준으로 주문량을 살펴본 결과, SP 지역에 주문량이 집중되는 것을 확인

지역별 배송 기간을 살펴본 결과, SP 지역을 제외한 다른 모든 지역의 평균 배송일이 10일 초과되는 것을 확인

주별 누적 주문량



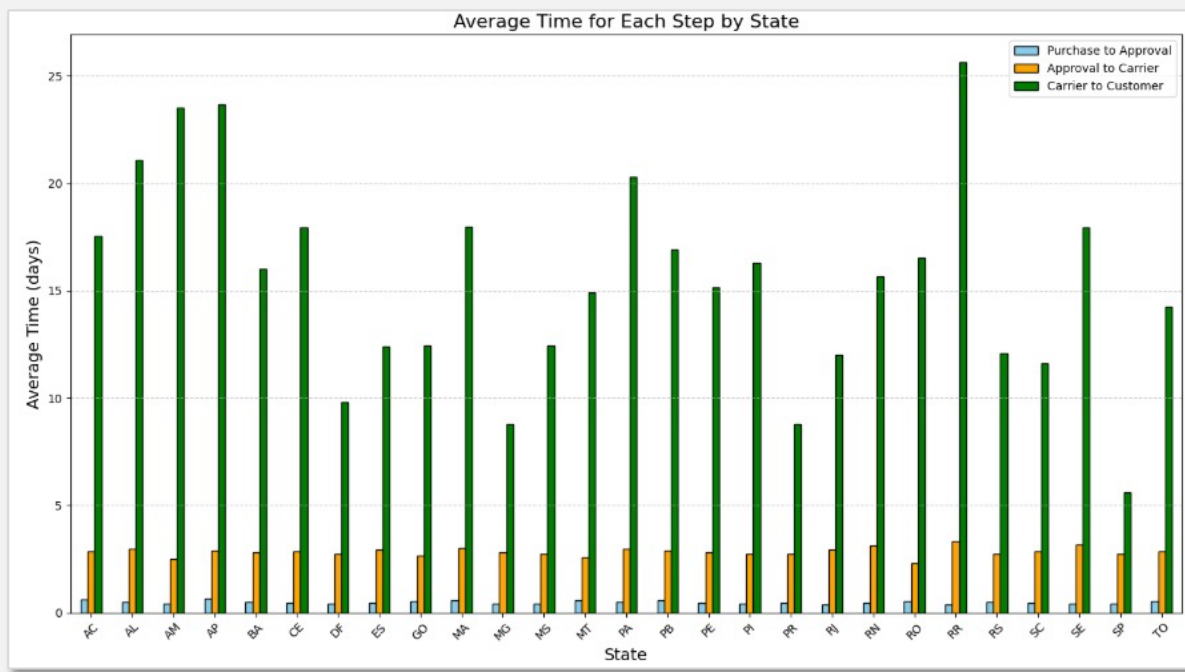
소비자 기준 주별 평균 배송 기간



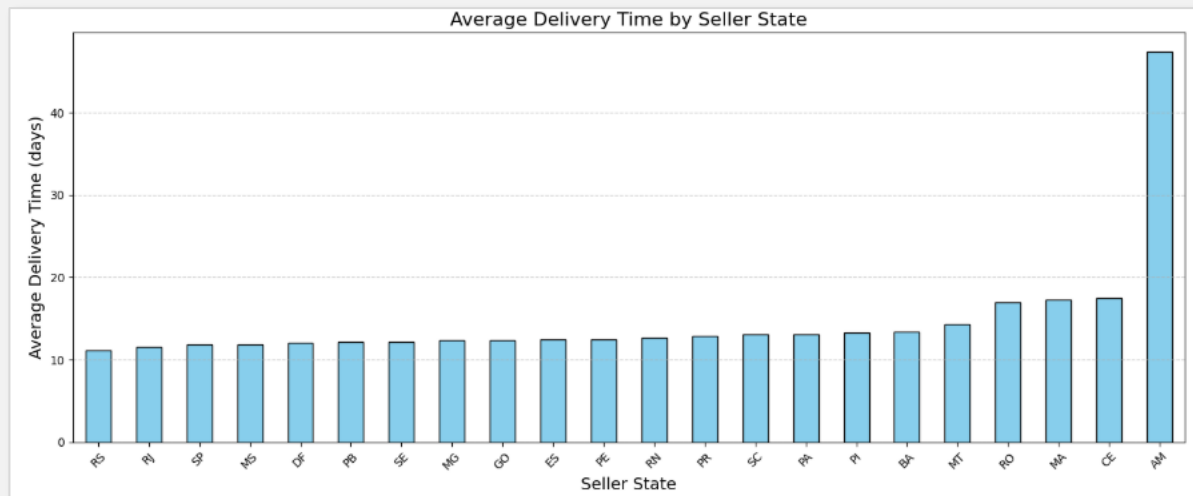
## 배송 문제 원인 분석

지역을 기준으로 배송의 각 단계별 평균 소요 시간을 확인한 결과, 택배사 – 배송 완료까지의 기간이 매우 상이함을 확인  
판매자 거주지역을 기준으로 평균 배송 소요기간을 확인한 결과, 모든 지역에서 비슷한 결과 확인

각 배송 단계에 따른 주별 소요 기간



판매자 기준 주별 평균 배송 기간

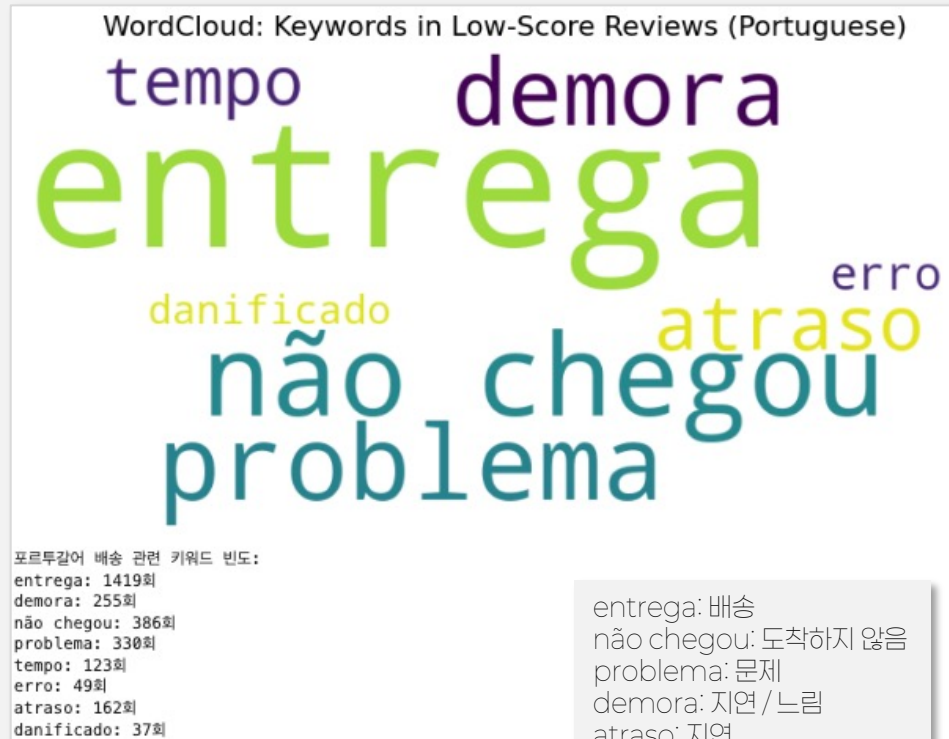


## 소비자 리뷰 분석

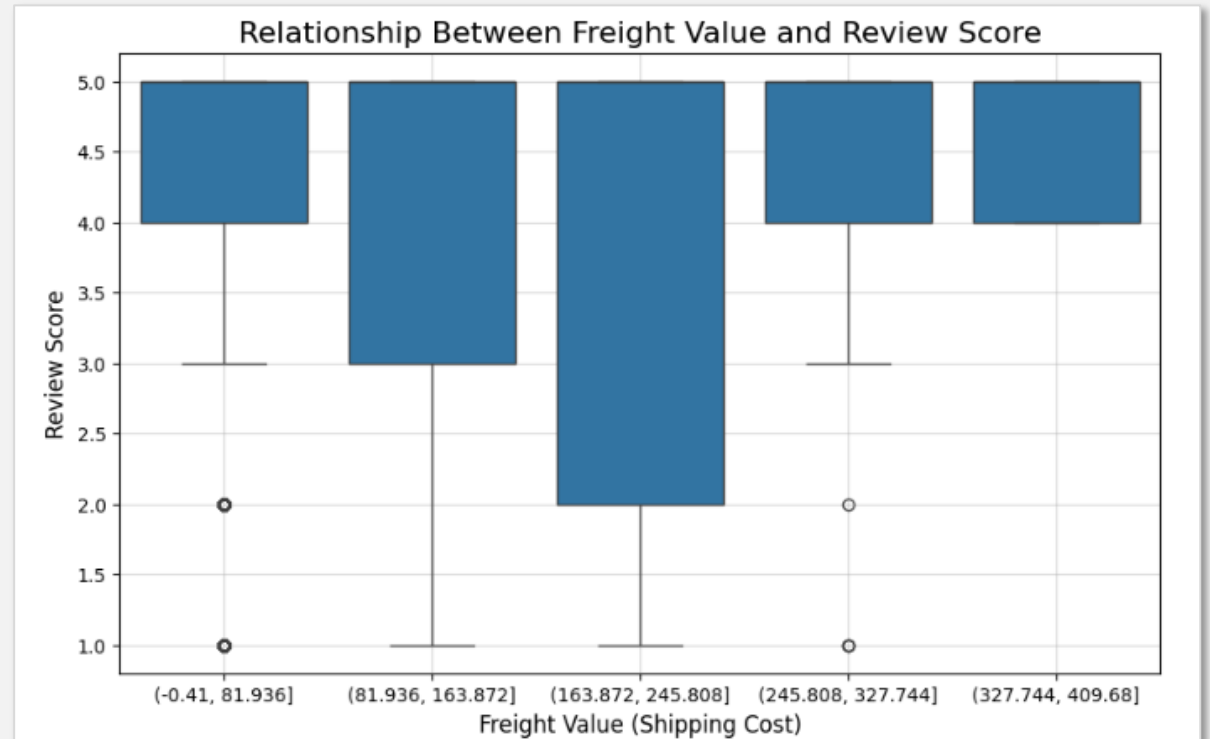
워드 클라우드 결과 배송과 관련된 문제를 제기하는 리뷰가 많았음

리뷰 점수를 확인한 결과, 배송 시간이 늘어날수록, 점수가 떨어지는 것을 확인

소비자 리뷰 분석 워드클라우드



배송 기간에 따른 리뷰 점수



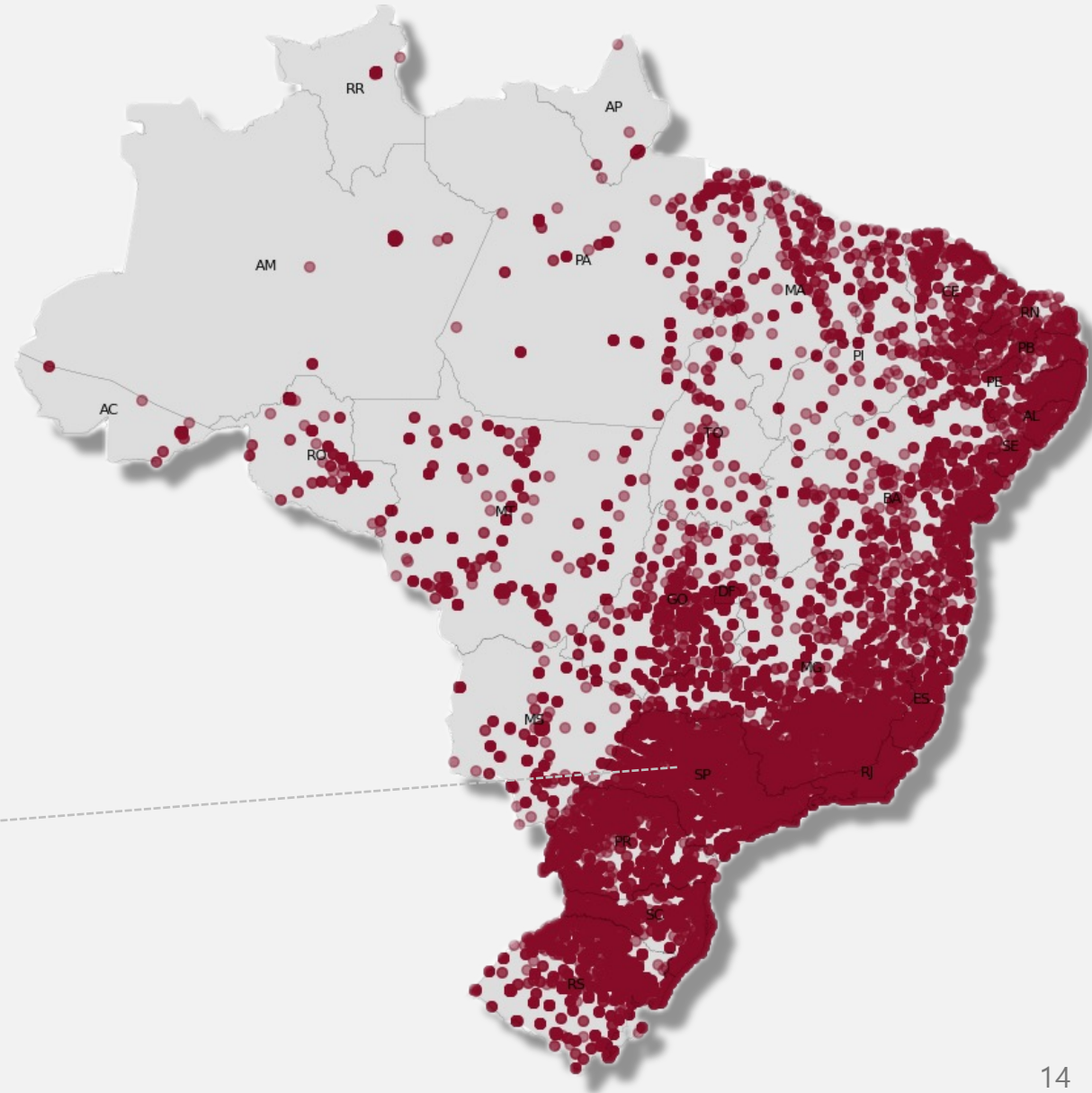
## 5. 해결 전략 제시

## 브라질 전 지역 배송망 구축

소비자 위치 기반 산점도 시각화 진행

: SP를 제외한 다른 지역의 배송 허브 구축을 통해 배송 기간 단축 기대

### São Paulo



# K-means를 활용한 배송 허브 위치 선정

각 지역의 위치 및 주문 수를 고려한 허브 위치 선정

배송의 신속성 및 수익성 고려

K-Medoids

Hierarchical  
Clustering

## K-means

데이터를 K개의 그룹으로  
나누는 방법으로써,  
랜덤으로 중심점을 정하고  
각 데이터를 가장 가까운 중심점에  
할당한 후 해당 과정 반복

$$J = \sum_{i=1}^K \sum_{x \in C_i} \|x - \mu_i\|^2$$

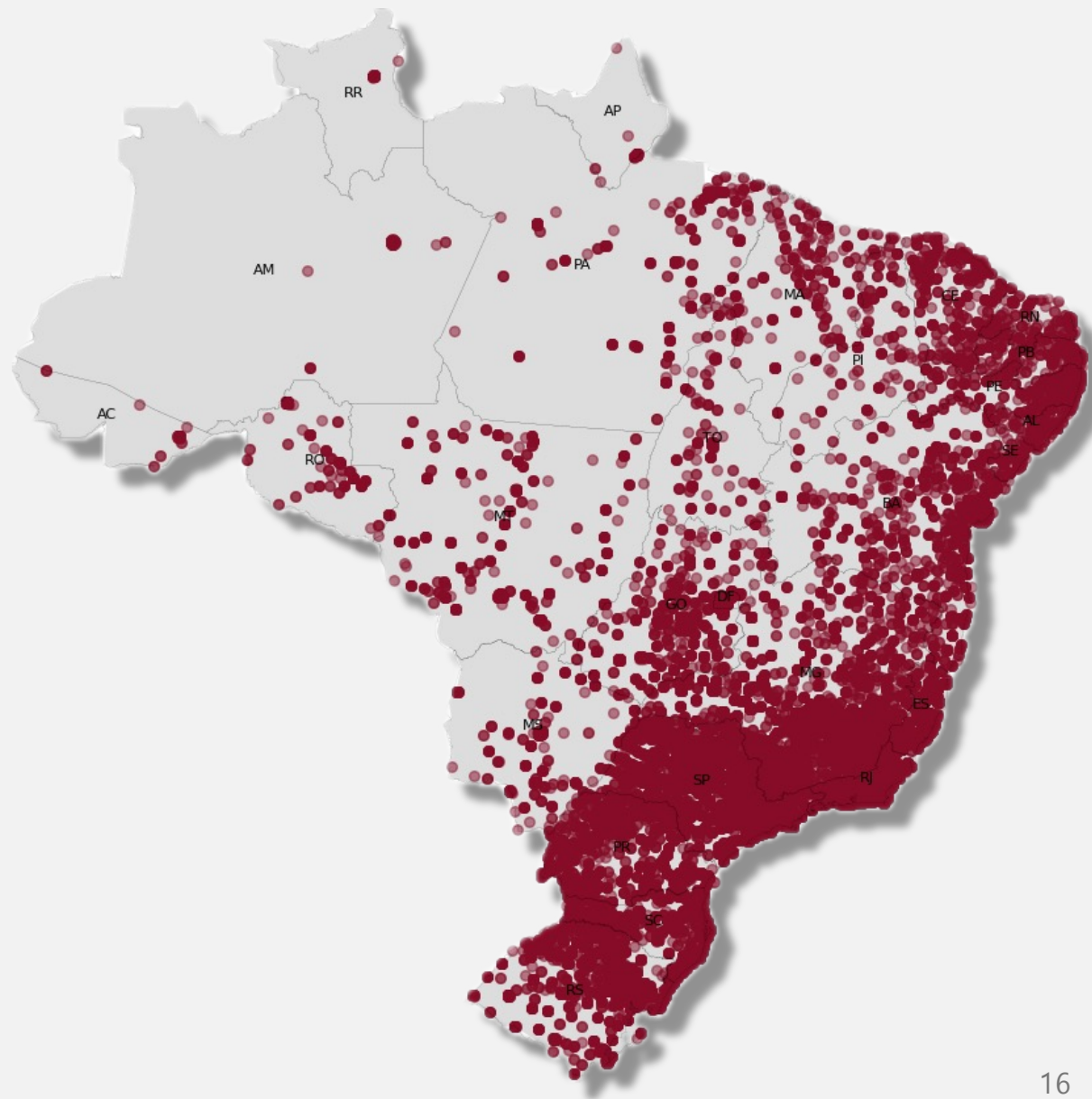
DBSCAN

OPTICS



## 서비스의 점유율이 높은 경우

전국으로 허브를 확대하는 것에 위험성이 낮다고 판단  
전국의 위치와 주문 수를 고려한 허브 위치 선정



허브를 추가한 지역

SC

RJ

PA

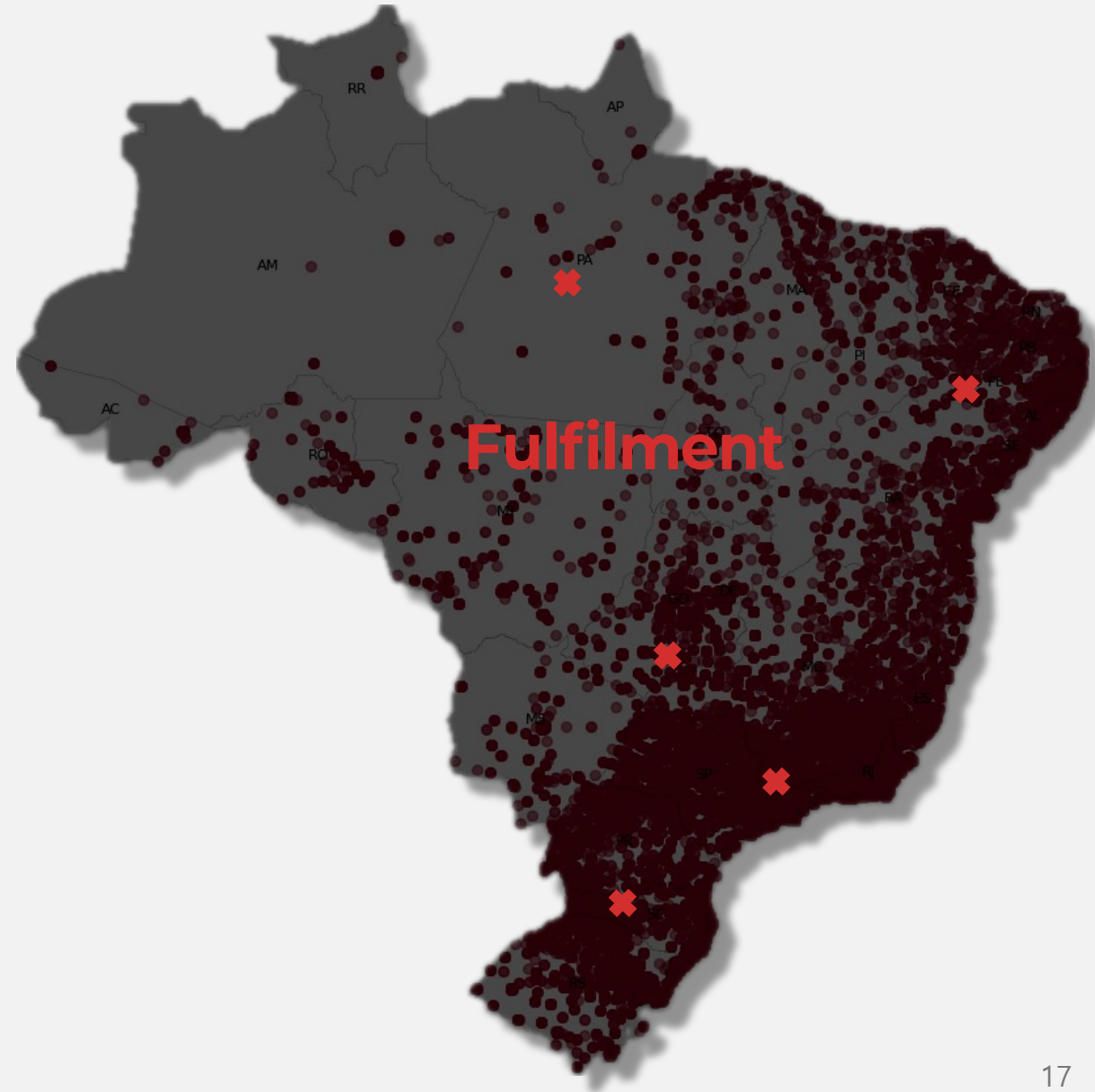
GO

PE



## 서비스의 점유율이 높은 경우

전국으로 허브를 확대하는 것에 위험성이 낮다고 판단  
전국의 위치와 주문 수를 고려한 허브 위치 선정



허브를 추가한 지역

SC

RJ

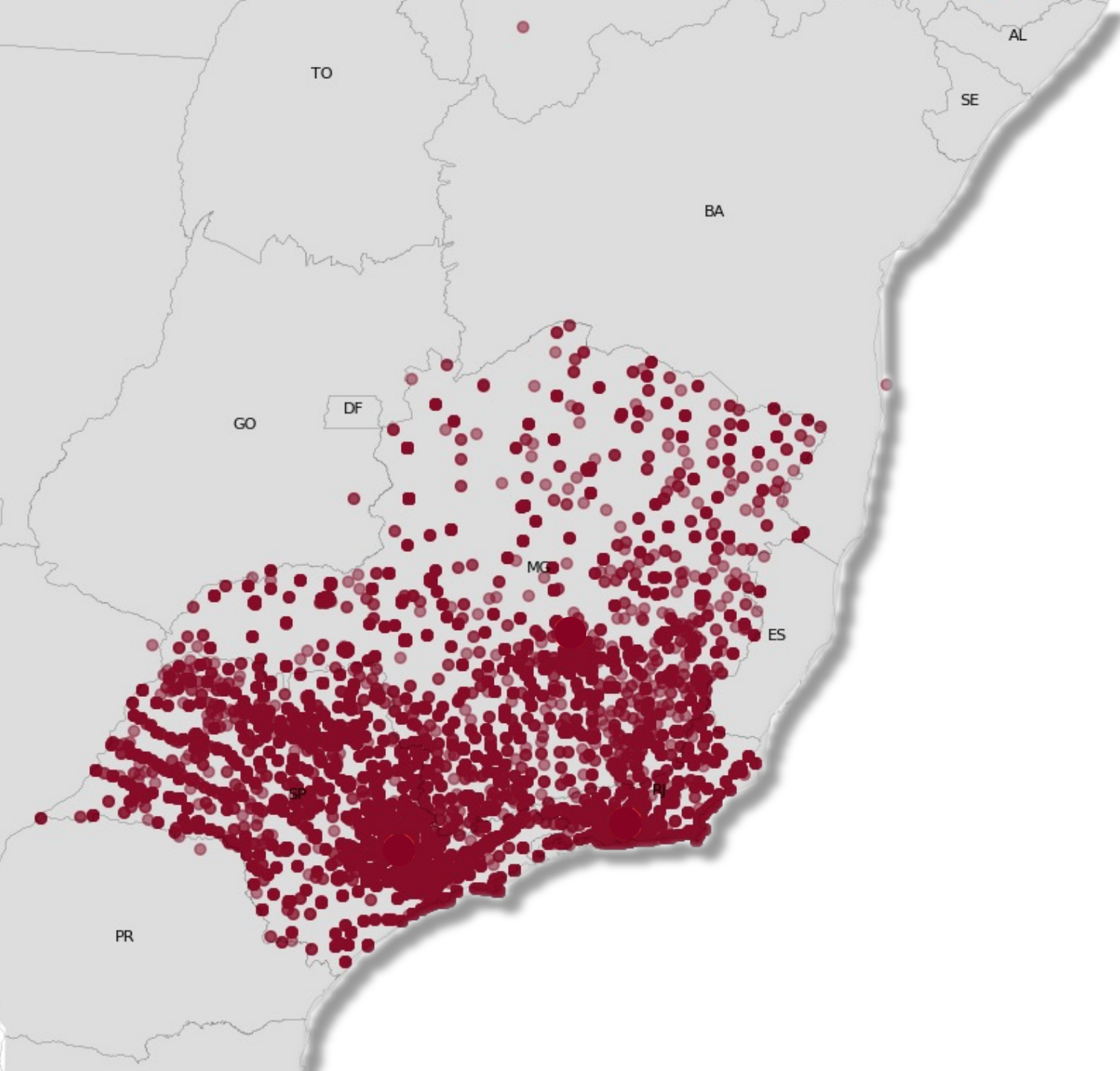
PA

GO

PE

## 서비스의 점유율이 낮은 경우

전국으로 허브를 확대하는 것에 위험성이 높다고 판단  
주문이 적정량 이루어지던 지역 2곳에 허브 확대 및 추이 관찰

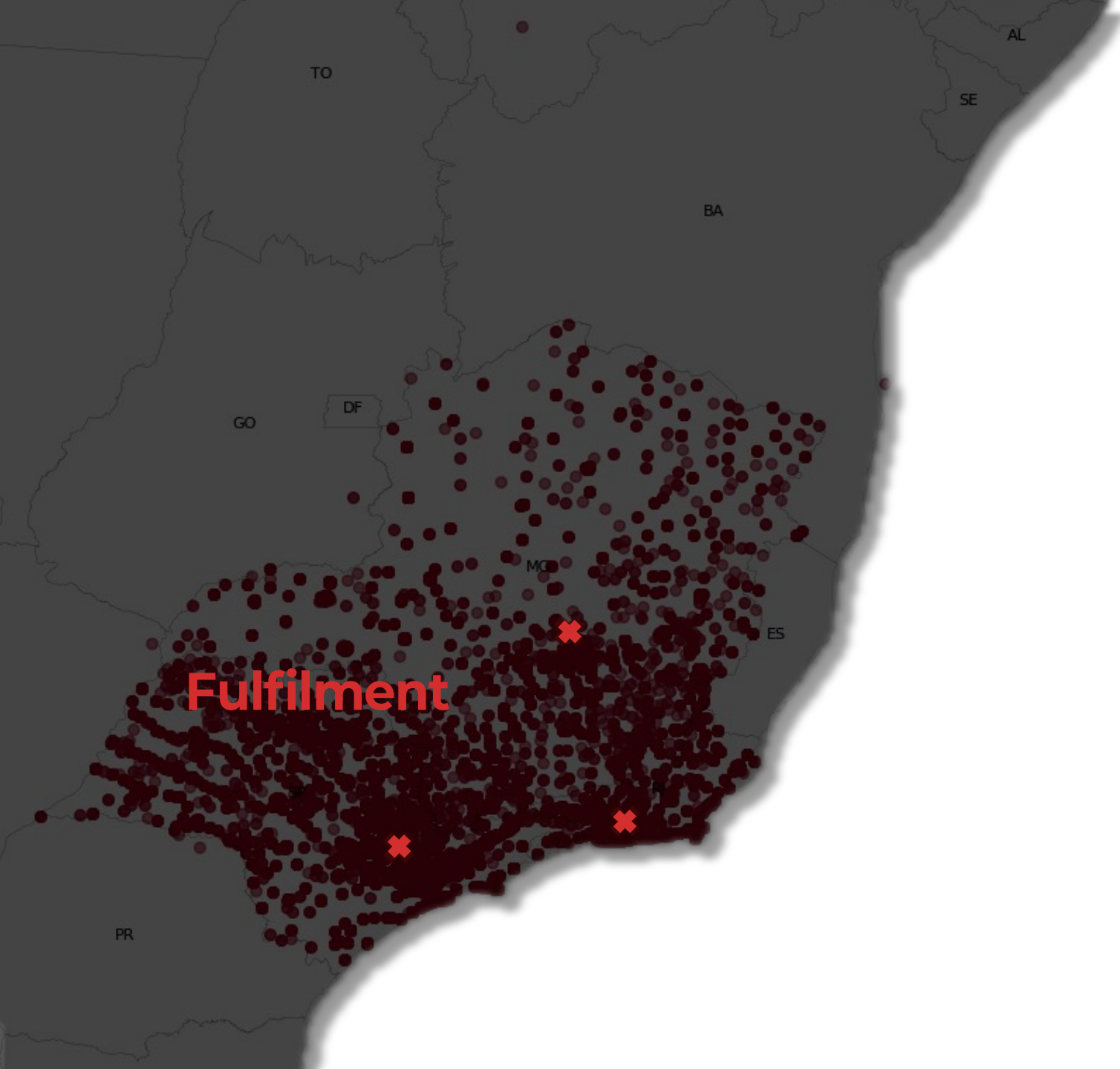


허브를 추가한 지역



## 서비스의 점유율이 낮은 경우

전국으로 허브를 확대하는 것에 위험성이 높다고 판단  
주문이 적정량 이루어지던 지역 2곳에 허브 확대 및 추이 관찰



허브를 추가한 지역



**감사합니다.**